DOCUMENT RESUME

ED 247 306 TM 840 488

AUTHOR Rose, Janet S.; And Others
TITLE Instructional Validity: Merging Curricular, Instructional and Test Development Issues.
PUB DATE Apr 84
NOTE 35p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Criterion Referenced Tests; *Curriculum Evaluation; Educational Objectives; Latent Trait Theory; Mathematics Achievement; *Test Construction; Test Interpretation; Test Use; *Test Validity
IDENTIFIERS *Curricular Validity; Rasch Model; *Test Curriculum Overlap

ABSTRACT

This paper proposes that instructional validity (i.e., the extent to which teachers actually teach the skills that are tested) should be part of the test development process. The development of a district-wide ninth-grade math exam for General Mathematics I (GMI) in the Charleston County School District, South Carolina, is used to illustrate how a study of the instructional process conducted during the early phases of test development can benefit the curriculum and strengthen the match between instruction and assessment. The GMI curriculum was divided into four nine-week grading periods. Teachers reported on the levels of instruction and student mastery of each stated objective. Reports on the students' progress were submitted quarterly and analyzed with regard to the Rasch model, which was used to generate item difficulties and fit statistics and to link the four sets of test forms. Analysis of the data found much variability in student mastery of course objectives. Comparisons among the four tests showed significant differences which have implications for the interpretation of the correlation coefficients. Topics for further research are outlined. Appended are the objectives covered by the GMI test. (Author/EGS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED247306

INSTRUCTIONAL VALIDITY: MERGING CURRICULAR, INSTRUCTIONAL AND TEST DEVELOPMENT ISSUES

Janet S. Rose
Charleston County (S.C.) School District

Joseph P. Ryan
College of Education
University of South Carolina

Ann T. Birdseye
Charleston County (S.C.) School District

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. S. Rose

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 840 488

Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1984.

Presentation: Instructional Validity: Merging Curricular, Instructional and Test Development Issues

Authors: Janet S. Rose, Charleston County (S.C.) School District
Joseph P. Ryan, University of South Carolina
Ann T. Birdseye, Charleston County (S.C.) School District

This paper proposes that instructional validity (i.e., the extent to which teachers actually teach skills tested) should be part of the test development process. The development of a districtwide ninth grade math exam is used to illustrate how a study of the instructional process conducted during the early phases of test development can benefit the curriculum and strengthen the match between instruction and assessment. Data collected from teachers' instructional records were analyzed in light of Rasch difficulty and fit statistics for each skill. Additional insights based upon teachers' responses on an end-of-year questionnaire were discussed.

INTRODUCTION

Ensuring a match between test items and curriculum content is a basic requirement for development of criterion-referenced tests. Test specifications, experts' judgmental reviews of test items, and item analysis procedures are commonly employed during the stages of test development to investigate and verify a test's content validity (the extent to which test content appropriately assesses specific skills of a particular curriculum). The courts have ruled, however, that demonstration of a test's content validity is not a sufficient condition for establishing that a test is "fundamentally fair in form and practice" (Debra P. v. Turlington, 1981). When tests are used for purposes which may deny constitutional guarantees of equal protection or due process — for example, when they are used to deny high school diplomas or grade promotion — the courts require that test users provide evidence that the test assesses skills actually taught in the classroom. School authorities cannot assume that state- or locally-mandated curricular objectives are taught uniformly in all classrooms. Rather, they must clearly demonstrate a test's "instructional validity" or the degree to which a test measures what has been taught to examinees.

Most attempts to establish that a test is instructionally valid for a particular curriculum have been post hoc. That is, the extent to which students received instruction in skills assessed by a test has been examined after the test has been developed and administered. A prime example is the SSAT-II, an examination required for high school graduation in the state of Florida and the subject of the Debra P. case. When use of this test was challenged, the court directed the state to produce evidence of its instructional validity. In situations where validity studies of a post hoc nature are performed, test users must resolve discrepancies between instructional content and test content by either revising their list of skills to match the test, or revising the test to

match instruction, and/or mandating and enforcing instruction of skills tested. None of these alternatives is particularly attractive once the test has been developed.

Rationale for the Study

Studying the congruence between test items and classroom instruction prior to the finalization of a test has certain advantages. First, problems of instructional validity can be avoided if data on the adequacy of instruction in skills assessed by test items are collected concurrently with the piloting of those items. Simultaneous collection of instructional data and pilot test information could result in saving time, energy, dollars and, some cases, litigation and postponement of the testing program. When a poor match is discovered in the early phases of test development, two options are available. Either objectives can be eliminated and matching test items deleted, or items can be retained and instructional staff can investigate reasons why teachers failed to teach the objectives in order to provide them with necessary assistance.

A second advantage of understanding the relationship between test items and instruction related to the items deals with the increased power that such an understanding would give to pyschometric analyses. Many new procedures for analyzing test items have been advanced in recent years. The rationale for these procedures has been based primarily in technical consideration, and few if any of these procedures have been advanced as providing new insights into the process of instruction. If a technical procedure is proposed as useful for educational measurement, then the procedure should show some sensitivity to known variation in instructional activities. The interpretation of psychometric analyses can be greatly enriched if the relationship between psychometric and instructional attributes of test items can be established.

A third value in examining instructional information as part of the test development process is that it may help to re-establish teachers as part of the assessment process. Many teachers and other educators have failing faith in testing programs because they see such programs as separate and distinct from the instructional program. Educators involved in curriculum and instruction may become more involved and supportive of testing if the testing community begins to recognize the value of instructional information in developing test design to assess the effects of instruction.

The testing program. Charleston County School District has integrated test development with an evaluation of locally-developed curriculum guides in elementary and middle school math and language arts and selected high school courses. In this paper the development of a districtwide subject area final exam for ninth grade General Mathematics I (GMI) will be used to illustrate how a study of instructional validity can merge and benefit both curriculum and test development efforts. Data collected from teachers' instructional records of GMI objectives were analyzed in light of (a) the percentage of items per objective classified as acceptable according to Rasch item fit statistics and (b) the Rasch average difficulty of acceptable items assessing the objective. It was hypothesized that (1) there would be a positive relationship between the extent to which an objective was taught by teachers and the percentage of fitting items remaining after the analysis. It was further hypothesized that (2) there would be a negative relationship between teachers' perceptions of the extent of students' mastery of an objective with the average difficulty of acceptable test items. (This hypothesis implies a postive relationship between students' mastery and item easiness.) Examples of how the results of the analyses assisted in revisions of the GMI curriculum and test are presented.

METHOD

GMI

The GMI curriculum is partitioned into four units corresponding to the four nine-weeks grading periods. Each unit has from 18 to 31 objectives which are tested. School district staff, ninth grade math teachers and external consultants prepared test specifications and items for each unit. Items were piloted at the end of each quarter during the 1982-83 school year. The Rasch model was used to generate item difficulties and fit statistics and to link the four sets of test forms. Fitting items were defined as those which had between-group fit statistics less than or equal to 3.0. A description of the technical issues regarding item analysis can be found in a report prepared by Educational Measurement Systems, Inc. (Ryan & Potter, 1983). The appendix contains the objectives covered by the GMI test.

Instructional Variables

Airasian and Madaus (1983) describe various procedures which have been applied to associate testing with instruction. Some compare test performance to elements of instructional materials, such as textbook and syllabi. Others analyze data on classroom instruction, either through direct observation, teacher interviews or teacher logs. (See Leinhardt (1981) for a review of these methods and for descriptions of studies in which they were used.)

Teacher self-report of instruction and student mastery, the method used here, was one component of an evaluation of the course curriculum guide that was conducted concurrently with the development of the GMI test. Teachers' estimates of students' opportunity to learn and their evaluations of students' level of mastery were gathered for GMI objectives. Fifty-one teachers were asked to assign one of the following descriptions of instruction and mastery to each objective as it applied to all students in their classes:

1. This objective was taught and mastered by most students (T/M).
2. Instruction was not necessary. Most students have mastered this objective (NT/M).
3. This objective was taught, but most students did not master it (T/NM).
4. Insufficient instructional time; objective not taught (NT/IT).
5. Objective not taught for some reason other than insufficient time (NT/O).

Teachers who failed to enter an instructional code for an objective were coded as "No Response" (NR).

Data from teachers' instructional records were collected at the end of each quarter for objectives for that quarter and for previous quarters. At the end of the school year, data from the four records were aggregated for each teachers and the percentage of teachers selecting each instructional code was calculated for each objective. Thirty-eight teachers (75%) returned completed records for the first and second quarter; 34 teachers (66%) returned forms at the end of the third quarter; and 32 teachers (63%) returned them at the end of the year. The total number of enrolled students per teacher ranged from 16 to 140 with a mean of 55.1 and a standard deviation of 29.8.

Analyses

Pearson-Product Moment correlations were generated for the two test descriptors — item difficulty and the percentage of fitting items per objective — and the percentage of teachers selecting each instructional code for each objective. In addition, instructional codes were collapsed in order to obtain and correlate the following instructional variables:

1. Mastered by students (M = T/M + NT/M).
2. Not mastered by students (NM = T/NM + NT/IT + NT/O).
3. Taught by teachers (T = T/M + T/NM).

4. Not taught by teachers (NT = NT/M + NT/IT + NT/O).

Initially, it was assumed that teachers who did not respond for a particular objective did not teach the objective. Therefore, teachers who did not respond were added in with NM and NT to create two additional variables: not mastered by most students plus no response (NMNO) and not taught by teachers plus no response (NTNO).

RESULTS

General Data Description

Tables 1-A through 1-D contain raw data for individual objectives from Tests 1 through 4, respectively, used to derive the correlation coefficients: the percentage of teachers selecting each code on the GMI instruction record (T/M, NT/M, T/NM, NT/IT, NT/O, NR); collapsed options (M, NM, NMNO, T, NT, NTNO); average difficulty estimates for retained items; and the percentage of items fitting the Rasch model which were retained in the item pool.

Table 2 presents descriptive statistics — mean, standard deviation, range (minimum and maximum) — for each variable. Overall, the average percentage of teachers teaching an objective was 60% (Tests 1-4), although the range (18-89) and standard deviation (20.6) indicate much variability in the degree to which different objectives were taught. The average percentage of teachers reporting that most of their students mastered a particular objective (54.3) was similar to the mean for non-mastery (42.8), as were the standard deviations and ranges. Again, much variability in student mastery of course objectives characterized GMI.

Comparisons among Tests

Comparisons among the four tests show significant differences (Tests 1 and 2 *v.* Test 3 *v.* Test 4).which have implications for the interpretation of the correlation coefficients. During the first two quarters the average

percentage of teachers teaching a particular objectives was approximately 74%; the mean for mastery was greater than 50%. Tests 3 and 4 show a trend for fewer objectives taught and a lower degree of student mastery. In addition, an average of 28.5% of the teachers did not enter instructional codes for Test 4 objectives. Though other reasons can be speculated, the two which were considered at the time were (a) the objectives were not taught and teachers did not want central staff to know this or (b) teachers were under pressure at the end of the school year and had too much paperwork to complete their records.

Statistics on the variables related to the GMI test items show that the tests increased in difficulty during the school year. (Note that the more positive difficulty estimates represent the more difficult items.) Also, overall, most test items fit the Rasch model and very few were deleted from the item bank. An average of 94% of the items written for a particular objective were retained. Percentages range from 75-100 for all except two objectives, objectives 206 and 208, for which only 22% of the items had acceptable fit statistics. The large percentage of retained items can be attributed to the procedures which were used to develop the GMI test. GMI teachers prepared the test specifications and helped write and revise the test items.

Correlations

Table 3 contains the correlation coefficients calculated for the instructional variables and the two test item variables. The correlations show the expected relationships between average item difficulty and extent of student mastery for Tests 1, 2 and 3, though only the correlations for Tests 1 and 2 reach statistical signficance. Thus, teachers reported that students tended to master objectives which had lower average item difficulties and failed to master objectives assessed by the more difficult items. Although the

correlations for NT/M and T/NM with average item difficulty were in the expected direction and significant for Test 3, those for the collapsed variables, i.e., M and NM, were not. We can speculate that differences in the distribution of data collected for the instructional codes, compared with Tests 1 and 2, may have contributed to the lower correlations. Comments made by GMI teachers on a supplemental questionnaire they completed at the end of the year and feedback from central office math staff suggested difficulties and problems regarding the probability objectives scheduled to be taught during this quarter. Apparently, teachers felt uncomfortable teaching these objectives and perhaps were unsure about student mastery. The findings for Test 4 were unusual — none of the expected correlations were found. However, a large proportion of teachers failed to complete their instructional records for many of the objectives. Given the relationships found for the previous tests, we cannot assume, at this point, that teachers did not teach and students did not master objectives omitted from teachers' instructional records. It appears that factors other than lack of instruction were responsible for the omissions. For this reason, analyses performed for variables that included "No Response," i.e., NR, NMNO and NTNO, cannot be interpreted.

Correlations between the percentage of fitting items and data collected via teachers' instructional records failed to show the strength that those for item difficulty did. The correlation between the extent to which teachers taught an objective and the percentage of fitting items was strong and positive for Test 2 only. The two objectives with 22% fitting items most likely contributed to the strong correlation for that test. A much lower percentage of teachers taught these objectives, and of those who did, more reported non-mastery than mastery.

In addition to investigating the two main hypotheses of this study, as second goal was to identify objectives and/or items to be revised or deleted.

To accomplish this, plots of the correlations between item difficulty and degree of student mastery and between the percentage of fitting items and extent of teacher instruction were drawn. These plots are illustrated in Figure 1 and 2, respectively. Three objectives — 326, 327 and 414 — circled in Figure 1 were of concern since they were relatively easy objectives which teachers indicated they did not have time to teach. Figure 2 identifies objectives 206 and 208 as objectives which just over half the teachers taught but had a small number of acceptable items.

DISCUSSION

The results of this study of instructional validity have beneficial outcomes. First, they demonstrate a true relationship between instructional events in the classroom and data derived from a field test designed for a particular course of study. Teachers' perceptions of the degree of student mastery were correlated with item difficulty. Also, if given some degree of variability in the quality of the items field-tested, the non-fitting items will tend to be those for which teachers provided less instruction.

The second and most important benefit is that instructional data can be used to address pertinent questions in the test development process, such as the following: Were items that did not fit the Rasch model of poor quality, or was the content assessed by the item not taught and/or mastered? Should the curriculum or the test be revised? Item fit data can be employed to answer these questions. For example, it appears that the large proportion of misfitting items for objectives 206 and 208 are not due to the poor quality of the test items but rather to low levels of instruction and/or poor instruction (i.e., relatively few teachers taught these objectives; of those who did more had students who they classified as non-masters). These two skills measured similar objectives — students' ability to convert square centimeters to square

millimeters and cubic centimeters to cubic millimeters — and were eliminated from the GMI curriculum by the math staff. In contrast, the one misfitting item for objective 201 (assessing metric estimation of line segments) was probably a poorly written item; 76% of the teachers reported that they taught this objective and 74% reported student mastery. Focusing on item difficulty and mastery, the data led to the elimination of the three outlying objectives circled in Figure 1: reading and making a bar graph (326); reading and making a line graph (327); solving word problems of a consumer and/or career nature involving coordinate graphing (414). It was interesting that the content of these three objectives was related. Since they were easy objectives characterized by low rates of instruction, they were eliminated from the curriculum to allow more time for teachers to teach the more difficult objectives.

A limitation of this study should not be overlooked. While item data were available for each student, only teacher level instructional and objective mastery data were collected. Obviously, results would be much more conclusive if teachers' reports of instruction and objective mastery could be collected at the student level and compared to each student's performance on items related to each objective. Unfortunately, such data collection efforts are very time consuming for teachers, but collecting data on instructional emphases at the student level or teacher level should be considered for the purpose of test development.

At a general level, the study suggests a useful line of research that has not been explored in great detail by those involved in educational measurement. The basic design for this line of research relates instructional variables to the psychometric attributes of test items written to assess the effects of instruction. This study operationalized instructional variables in terms of teacher self-reports, and other procedures for operationalizing

instructional variables need to be explored (e.g., direct classroom observation). Furthermore, this study examined item difficulty and item fit for the one-parameter latent trait model. Future studies might explore the relationship between instructional variables and item discrimination, in both the classical and latent trait sense, and the tendency of items to elicit guessing. The interpretation of our psychometric analyses might be enhanced if we could understand whether or to what extent the psychometric attributes of test items are revealed in the instructional activities of classroom teachers.

References

Airasian, P. W., & Madaus, G. F. (1983). Linking testing and instruction: Policy issues. Journal of Educational Measurement, 20, 103-118.

Debra P. V. Turlington, 644 F.2d 397 (5th Cir. 1981).

Leinhardt, G. (1983). Overlap: Testing whether it is taught. In G. Madaus (Ed.), The courts, validity, and minimum competency testing (pp. 151-168). Boston: Kluwer-Nijhoff Publishing.

Ryan, J. P., & Potter, D. C. (1983). Procedures used to create the ninth grade general mathematics item bank for the Charleston County School District. Educational Measurement Systems, Inc.

Table 1-A

Percentage of Teachers Selecting Response Options on the General Mathematics I Instructional Record Test 1 Objectives

Obj	T/M	NT/M	T/NM	NT/IT	NT/O	NR	M	NM	NMNO	T	NT	NTNO	Avg. Diff.	% Items
103	66	5	8	3	8	11	71	19	30	74	16	27	-1.5909	100
105	61	5	5	15	5	8	66	25	33	66	25	33	-0.1801	100
108	82	16	3	0	0	0	98	3	3	85	16	16	-2.5923	100
114	68	5	21	3	0	3	73	24	27	89	8	11	-0.6653	100
115	71	3	13	0	3	11	74	16	27	84	6	17	0.7091	100
117	71	3	11	5	0	11	74	16	27	82	8	19	-0.8167	100
118	53	5	29	11	0	3	58	40	43	82	16	19	-0.9353	100
119	66	3	13	11	0	8	69	24	32	79	14	22	-0.1993	100
120	82	8	0	8	0	3	90	8	11	82	16	19	-2.8723	89
121	79	5	5	8	0	3	84	13	16	84	13	16	-2.6180	89
122	61	0	18	13	0	8	61	31	39	79	13	21	-1.3739	100
123	68	3	8	13	3	5	71	24	29	76	19	24	-1.8790	100
125	58	3	13	18	3	5	61	34	39	71	24	29	0.3001	100
126	50	0	26	16	3	5	50	45	50	76	19	24	-0.2582	100
128	42	3	21	21	3	11	45	45	56	63	27	38	-0.6099	100
129	18	0	18	42	5	16	18	65	81	36	47	63	0.7376	89
130	24	0	29	37	3	8	24	69	77	53	40	48	0.4576	100
131	29	0	13	42	8	8	29	63	71	42	50	58	0.3909	100

Notes: Column headings are as follows:
Obj = Objective;
T/M = Taught and mastered;
NT/M = Not taught; already mastered by students;
T/NM = Taught, but not mastered;
NT/IT = Not taught; lack of instructional time;
NT/O = Not taught; other reason;
NR = No response;
M = Mastered by students (T/M + NT/M);
NM = Not mastered by students (T/NM + NT/IT + NT/O);

Table 1-A
Test 1 (cont'd)

Notes: NMNC = NM + NR;
T = Taught by teachers (T/M + T/NM);
NT = Not taught by teachers (NT/M + NT/IT + NT/O);
NTNO = NT + NR;
Avg. Diff. = Average difficulty of items in bank;
% Items = Percentage of items piloted which fit the Rasch model.

Table 1-B

Percentage of Teachers Selecting Response Options on the
General Mathematics I Instructional Record
Test 2 Objectives

Obj	T/M	NT/M	T/NM	NT/IT	NT/O	NR	M	NM	NMNO	T	NT	NTNO	Avg. Diff.	% Items
201	71	3	5	11	3	8	74	19	27	76	17	25	-1.0771	89
202	61	0	18	13	3	5	61	34	39	79	16	21	-1.1627	100
203	42	0	29	16	5	8	42	50	58	71	21	29	1.0622	100
204	47	0	29	13	3	8	47	45	53	76	16	24	1.4039	80
205	68	0	11	13	3	5	68	27	32	79	16	21	-0.9539	100
206	24	0	32	24	8	13	24	64	77	56	32	45	0.4022	22
207	58	0	8	18	8	8	58	34	42	66	26	34	-0.0729	89
208	24	0	32	26	8	11	24	66	77	56	34	45	0.7232	22
209	50	0	29	13	3	5	50	45	50	79	16	21	-0.1922	100
210	47	0	32	13	3	5	47	48	53	79	16	21	0.8122	78
211	50	0	29	13	3	5	50	45	50	79	16	21	-1.0605	100
212	37	0	37	16	3	8	37	56	64	74	19	27	0.9228	78
213	32	0	32	26	3	8	32	61	69	64	29	37	0.0739	67
214	32	0	27	21	0	11	32	58	69	69	21	32	0.0827	100
215	68	5	16	11	0	0	73	27	27	84	16	16	-1.6010	100
217	71	0	16	11	0	3	71	27	30	87	11	14	-1.4537	100
218	84	0	3	11	0	3	84	14	17	87	11	14	-1.2523	89
219	55	0	24	11	0	11	55	35	46	79	11	22	-0.1949	100
220	47	0	32	11	0	11	47	43	54	79	11	22	0.3615	100
222	79	3	3	13	0	3	82	16	19	82	16	19	-0.3820	80
223	74	0	8	13	0	5	74	21	26	82	13	18	-0.6745	80
224	74	3	5	13	0	5	77	18	23	79	16	21	-0.4895	100

(continued on next page)

Table 1-B
Test 2 (cont'd)

Obj	T/M	NT/M	T/NM	NT/IT	NT/O	NR	M	NM	NMNO	T	NT	NTNO	Avg. Diff.	% Items
227	61	3	18	16	0	3	64	34	37	79	19	22	0.2939	100
228	39	0	32	21	0	8	39	53	61	71	21	29	0.6010	89
230	63	0	16	18	0	3	63	34	37	79	18	21	-0.8869	100
233	63	3	13	16	0	5	66	29	34	76	19	24	-0.0439	80
234	53	3	21	18	0	5	56	39	44	74	21	26	0.0082	100
235	34	3	26	26	0	10	37	52	62	60	29	39	0.2389	100

Notes: Column headings are as follows:
Obj = Objective;
T/M = Taught and mastered;
NT/M = Not taught; already mastered by students;
T/NM = Taught, but not mastered;
NT/IT = Not taught; lack of instructional time;
NT/O = Not taught; other reason;
NR = No response;
M = Mastered by students (T/M + NT/M);
NM = Not mastered by students (T/NM + NT/IT + NT/O);
NMNO = NM + NR;
T = Taught by teachers (T/M + T/NM);
NT = Not taught by teachers (NT/M + NT/IT + NT/O);
NTNO = NT + NR;
Avg. Diff. = Average difficulty of items in bank;
% Items = Percentage of items piloted which fit the Rasch model.

Table 1-C

Percentage of Teachers Selecting Response Options on the General Mathematics I Instructional Record Test 3 Objectives

Obj	T/M	NT/M	T/NM	NT/IT	NT/O	NR	M	NM	NMNO	T	NT	NTNO	Avg. Diff.	% Items
302	29	0	21	24	18	9	29	63	72	50	42	51	1.3291	100
305	21	0	24	24	21	12	21	69	81	45	45	57	-0.0951	100
306	24	0	26	24	18	9	24	68	77	50	42	51	1.2922	100
309	74	0	12	12	0	3	74	24	27	86	12	15	-0.4591	100
310	24	0	29	32	6	9	24	67	76	53	38	47	-0.6163	100
311	32	0	18	38	6	6	32	62	68	50	44	50	-0.5709	88
312	50	0	6	35	6	3	50	47	50	56	41	44	-0.6011	100
313	35	0	12	35	9	9	35	56	65	47	44	53	-0.9043	100
314	59	0	12	15	6	9	59	33	42	71	21	30	-1.0126	100
315	50	0	13	13	5	19	50	31	50	63	18	37	-0.7807	100
316	47	0	21	18	6	9	47	45	54	68	24	33	0.3496	100
317	47	0	24	15	6	9	47	45	54	71	21	30	-0.0169	89
318	82	0	6	12	0	0	82	18	18	88	12	12	-0.1593	89
319	74	0	12	12	0	3	74	24	27	86	12	15	0.3712	100
320	65	0	18	15	0	3	65	33	36	83	15	18	0.3388	100
321	44	0	38	15	0	3	44	53	56	82	15	18	1.1108	100
322	35	0	47	15	0	3	35	62	65	82	15	18	0.8533	100
323	29	0	35	26	3	6	29	64	70	64	29	35	0.3651	100
324	47	0	0	38	3	12	47	41	53	47	41	53	0.6942	100
325	47	0	0	38	3	12	47	41	53	47	41	53	0.1608	89
326	44	3	3	38	3	9	47	44	53	47	44	53	-2.1466	100
327	44	3	3	38	3	9	47	44	53	47	44	53	-1.8057	100

(continued on next page)

Table 1-C
Test 3 (cont'd)

Obj	T/M	NT/M	T/NM	NT/IT	NT/O	NR	M	NM	NMNO	T	NT	NTNO	Avg. Diff.	% Items
328	44	0	3	38	3	12	44	44	56	47	41	53	-1.4267	89
329	15	0	12	47	9	18	15	68	86	27	56	74	0.4803	100
330	35	0	3	44	9	9	35	56	65	38	53	62	-0.6254	100
331	32	0	6	41	9	12	32	56	68	38	50	62	-0.0368	100
332	29	0	6	44	9	12	29	59	71	35	53	65	0.6958	89
333	26	0	12	50	3	9	26	65	74	38	53	62	0.1655	89
334	21	0	15	50	3	12	21	68	80	36	53	65	1.6493	75
336	24	0	6	50	6	15	24	62	77	30	56	71	-0.3895	100
337	9	0	9	62	6	15	9	77	92	18	68	83	0.0004	75

Notes: Column headings are as follows:
Obj = Objective;
T/M = Taught and mastered;
NT/M = Not taught; already mastered by students;
T/NM = Taught, but not mastered;
NT/IT = Not taught; lack of instructional time;
NT/O = Not taught; other reason;
NR = No response;
M = Mastered by students (T/M + NT/M);
NM = Not mastered by students (T/NM + NT/IT + NT/O);
NMNO = NM + NR;
T = Taught by teachers (T/M + T/NM);
NT = Not taught by teachers (NT/M + NT/IT + NT/O);
NTNO = NT + NR;
Avg. Diff. = Average difficulty of items in bank;
% Items = Percentage of items piloted which fit the Rasch model.

Table 1-D

Percentage of Teachers Selecting Response Options on the
General Mathematics I Instructional Record
Test 4 Objectives

Obj	T/M	NT/M	T/NM	NT/IT	NT/O	NR	M	NM	NMNO	T	NT	NTNO	Avg. Diff.	% Items
402	38	0	16	19	0	28	38	35	63	54	19	47	0.3244	100
403	38	0	16	16	0	31	38	32	63	54	16	47	-0.1079	100
404	34	0	19	16	0	31	34	35	66	53	16	47	0.2583	100
405	28	0	9	28	0	34	28	37	71	37	28	62	0.5999	100
406	25	0	13	28	0	34	25	41	75	38	28	62	0.5045	100
407	25	0	13	28	3	31	25	44	75	38	31	62	-0.2436	100
411	22	0	6	41	0	31	22	47	78	28	41	72	-0.2659	89
412	16	0	9	41	0	34	16	50	84	25	41	75	0.4094	100
413	13	0	13	41	0	34	13	54	88	26	41	75	0.6310	89
414	6	0	19	41	3	31	6	63	94	25	44	75	-1.3569	100
415	53	0	9	16	0	22	53	25	47	62	16	38	0.4305	100
416	47	0	13	19	0	22	47	32	54	60	19	41	0.7178	100
417	50	0	9	19	0	22	50	28	50	59	19	41	0.0876	100
418	44	0	16	19	0	22	44	35	57	60	19	41	1.5398	89
419	34	0	16	28	0	22	34	44	66	50	28	50	1.5948	89
420	25	0	9	44	0	22	25	53	75	34	44	66	0.3384	100
424	19	0	6	44	0	31	19	50	81	25	44	75	1.0521	100
425	13	0	9	47	0	31	13	56	87	22	47	78	1.3881	100
426	9	0	13	47	0	31	9	60	91	22	47	78	1.0659	100
427	9	0	13	47	0	31	9	60	91	22	47	78	1.1811	100
428	16	0	9	50	0	25	16	59	84	25	50	75	0.8726	100
430	22	0	9	41	0	28	22	50	78	31	41	69	0.5518	100

(continued on next page)

Table 1-D
Test 4 (cont'd)

Notes: Column headings are as follows:
Obj = Objective;
T/M = Taught and mastered;
NT/M = Not taught; already mastered by students;
T/NM = Taught, but not mastered;
NT/IT = Not taught; lack of instructional time;
NT/O = Not taught; other reason;
NR = No response;
M = Mastered by students (T/M + NT/M);
NM = Not mastered by students (T/NM + NT/IT + NT/O);
NMNO + NM + NR;
T = Taught by teachers (T/M + T/NM);
NT = Not taught by teachers (NT/M + NT/IT + NT/O);
NTNO = NT + NR;
Avg. Diff. = Average difficulty of items in bank;
% Items = Percentage of items piloted which fit the Rasch model.

Table 2

Descriptive Statistics for Variables Used in Analyses

Variable	TEST 1 (18 objectives)			TEST 2 (16 objectives)			TEST 3 (31 objectives)			TEST 4 (22 objectives)			TESTS 1-4 (99 objectives)		
	Mean	S.D.	Range Min/Max	Mean	S.D.	Range Min/Max	Mean	S.D.	Range Min/Max	Mean	S.D.	Range Min/Max	Mean	S.D.	Range Min/Max
Taught and mastered (1)	58.3	19.1	18/82	53.9	16.8	24/84	39.9	17.7	9/82	26.6	13.9	6/53	44.3	20.4	6/84
Not taught; already mastered (2)	3.7	3.8	0/16	0.9	1.5	0/5	0.2	0.7	0/3	0	0	-	1.0	2.3	0/16
Taught; not mastered (3)	14.1	8.7	0/29	21.2	11.0	3/37	14.6	11.6	0/47	12.0	3.9	6/19	15.8	10.2	0/47
Not taught; lack of time (4)	14.7	13.2	0/42	15.9	4.9	11/26	30.9	14.3	12/62	32.7	12.3	16/50	24.1	14.1	0/62
Not taught; other reason (5)	2.4	2.7	0/8	2.0	2.6	0/8	5.8	5.3	0/21	0.3	0.9	0/3	2.9	4.0	0/21
No response (6)	7.1	4.0	0/16	6.5	3.2	0/13	9.0	4.6	0/19	28.5	4.6	22/34	12.3	9.7	0/34
Mastered by students (1+2)	62.0	21.9	18/98	54.8	17.4	24/84	40.1	17.7	9/82	26.6	13.9	6/53	45.3	21.6	6/98
Not mastered by students (3+4+5)	31.3	19.6	3/69	39.1	15.1	14/66	51.3	15.9	18/77	45.0	11.4	25/63	42.8	16.8	3/77
Not mastered by students plus (6)	38.4	21.6	3/81	45.6	17.4	17/77	60.3	17.8	18/92	73.5	13.8	47/94	55.1	21.5	3/94
Taught by teacher (1+3)	72.4	15.1	36/89	75.0	8.3	56/87	54.5	19.2	18/88	38.6	14.8	22/62	60.0	20.6	18/89
Not taught by teacher (2+4+5)	20.9	12.8	6/50	18.8	6.2	11/34	36.9	16.0	12/68	33.0	12.4	16/50	28.0	14.6	6/68
Not taught by teacher plus (6)	28.0	14.8	11/63	25.4	8.3	14/45	45.9	19.2	12/83	61.5	14.6	38/78	40.3	20.5	11/83
Item difficulty	-0.7776	1.1603	-2.8723/ 0.7376	-0.1610	0.8100	-1.6010/ 1.4039	-0.0577	0.8942	-2.1466/ 1.6493	0.5281	0.6766	-1.3569/ 1.5948	-0.0881	0.9649	-2.8723/ 1.6493
Percentage of fitting items	98.2	4.2	89/100	87.3	20.9	22/100	95.9	7.3	75/100	98.0	4.3	89/100	94.3	12.8	22/100

Table 3

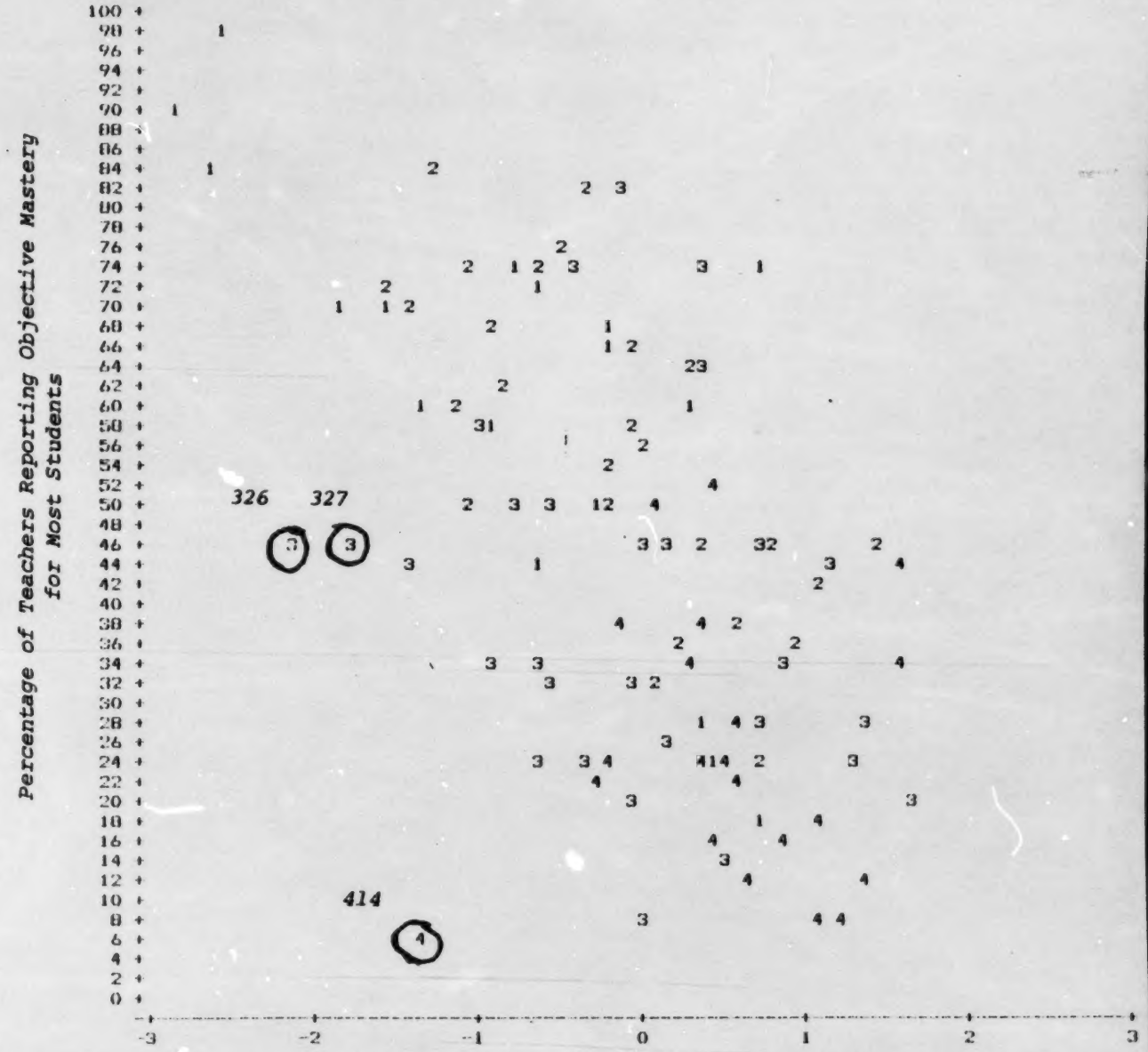
Correlation Coefficients for GMI Quarters 1 - 4

Instructional Code	Item Difficulty					Percentage of Fitting Items				
	Q1	Q2	Q3	Q4	Q1-Q4	Q1	Q2	Q3	Q4	Q1-Q4
T/M	-.71***	-.70***	-.23	.04	-.55***	-.03	.48**	.25	-.06	.11
NT/M	-.66**	-.24	-.57***	0	-.55***	-.07	.21	.15	0	.09
T/NM	.56*	.65***	.46**	-.17	.36***	.34	-.26	.19	-.09	-.19
NT/IT	.57*	.45*	-.09	.14	.33***	-.16	-.54**	-.47**	.02	-.06
NT/O	.44	.33	.17	-.63**	.08	.13	-.64***	.15	.15	-.13
M	-.74***	-.70***	-.25	.04	-.58***	-.04	.48**	.25	-.06	.11
NM	.69**	.68***	.32	.04	.52***	.06	-.47*	-.24	0	-.20*
NMNO	.74***	.70***	.26	-.04	.58***	.05	-.49**	-.25	.05	-.12
T	-.58*	-.56**	.06	-.01	-.37***	.15	.63***	.34	-.08	.01
NT	.47*	.44*	-.05	.09	.26**	-.16	-.65***	-.36**	.03	-.08
NTNO	.58*	.56**	-.06	.01	.37***	-.15	-.66***	-.35	.07	-.02

*$p < .05$
**$p < .01$
***$p < .001$

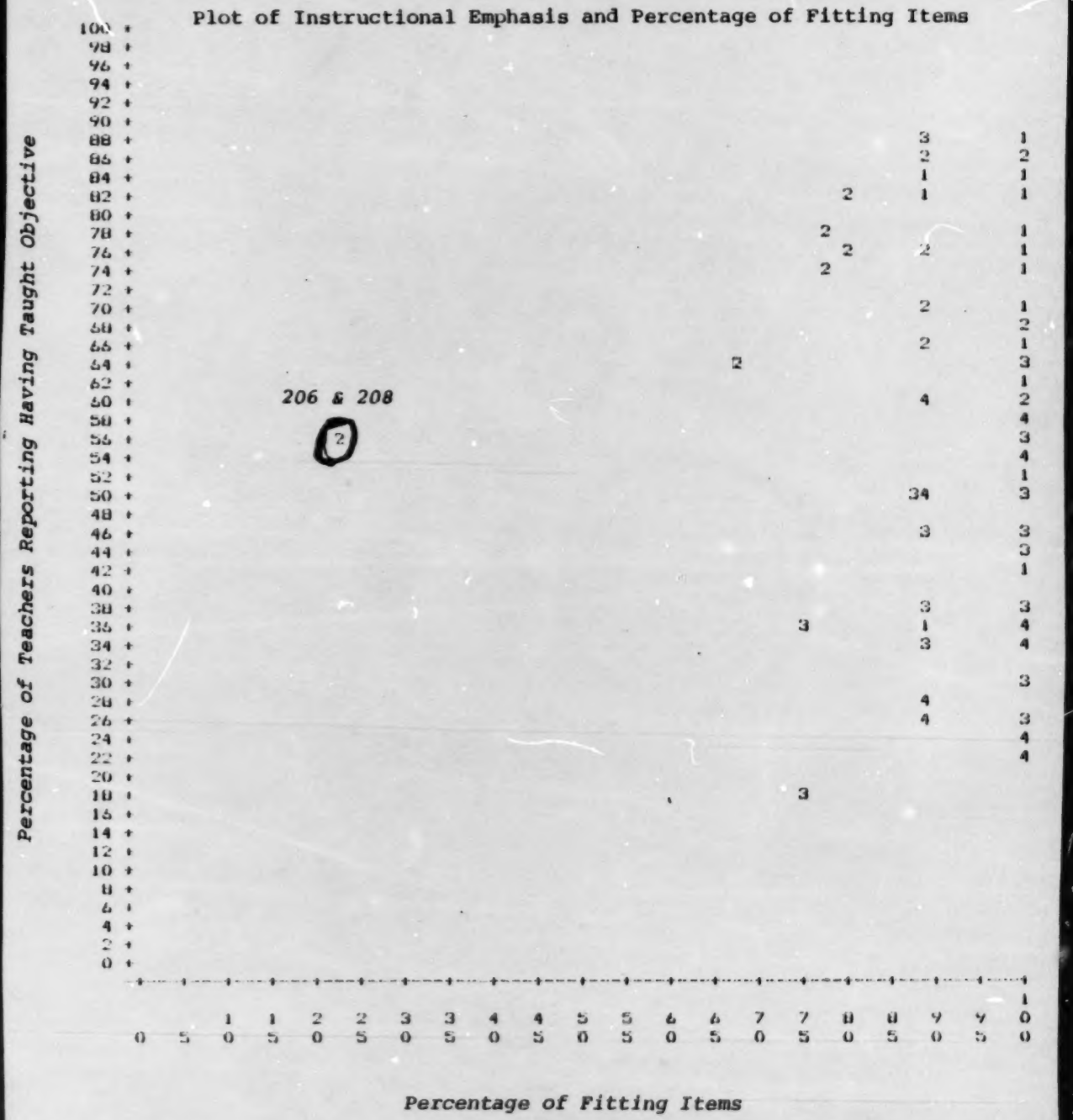
Figure 1

Plot of Student Mastery and Item Difficulty
(Symbol is Test)



BEST COPY AVAILABLE

Figure 2


Plot of Instructional Emphasis and Percentage of Fitting Items
Percentage of Teachers Reporting Having Taught Objective
206 & 208
Percentage of Fitting Items

Ninth Grade General Mathematics I Objectives

Assessed by the GMI Area Exam

Level 1

Code	Objective: "The student will be able to. . ."
103	solve equations in one variable using addition and subtraction.
105	read and write whole numbers and standard numerals for powers of 10 and vice versa.
108	multiply whole numbers with a 2-digit multiplier.
114	divide whole numbers with a 3-digit divisor.
115	use order of operations for computation involving addition, subtraction, multiplication and division.
117	identify the place value of a decimal through the millionth place.
118	read and write the word name of a mixed decimal fraction or vice versa.
119	compare decimals by selecting the correct inequality or equality sign that indicates their order.
120	add decimals.
121	subtract decimals.
122	solve word problems of a consumer and/or career nature involving addition and subtraction of decimals.
123	multiply decimals.
125	multiply or divide a decimal by 10, 100 or 1,000 (moving the decimal point to the left or right).
126	divide a decimal by a decimal with a one, two or three place decimal in the divisor.
128	round decimal quotients to the nearest tenth, hundredth, or thousandth.
129	write a standard numeral for a number expressed in scientific notation.
130	solve word problems of a consumer and/or career nature involving multiplication and division of decimals.
131	solve addition and subtraction problems involving different times.

Ninth Grade General Mathematics I Objectives
Assessed by the GMI Area Exam

Level 2

Code	Objective: The student will be able to. . ."
201	estimate and measure the length of a segment to the nearest centimeter and/or millimeter.
202	select the appropriate linear measure using the metric units - kilometers, meters and centimeters.
203	Add two measurements expressed in yards and inches or in feet when given a problem requiring regrouping.
204	make conversions of linear measure using the metric units kilometers, meters, and centimeters.
205	find the area of figures by counting square centimeters.
206	convert square centimeters into square millimeters.
207	find the volume of figures by counting cubic centimeters.
208	convert cubic centimeters to cubic millimeters.
209	select the appropriate metric measure to measure the capacity of a given container (liter, milliliter).
210	make conversions from liter to milliliter and vice versa.
211	select the appropriate measure to measure the weight of a given object (metric or standard).
212	make conversions between the following metric units of mass: kilogram, gram, milligram.
213	when given labeled pictures of thermometers in Celsius, select the appropriate metric unit of temperatures from descriptive phrases dealing with temperatures such as warm bread, snowball, etc.
214	solve word problems of consumer and/or career nature dealing with metric units of measure (length, mass, capacity and temperature).
215	write a fraction or a mixed number for shaded regions and/or number lines.
217	find equal ratios by multiplication.
218	change improper fractions to mixed numbers or whole numbers and vice versa.
219	write a fraction as a terminating decimal.

Ninth Grade General Mathematics I Objectives
Assessed by the GMI Area Exam

Level 2 (continued)

Code	Objectives: The student will be able to. . ."
220	write a faction as a nonterminating decimal and round to the nearest thousandth.
222	multiply a proper fraction by a whole number.
223	multiply three or more proper fractions.
224	write the reciprocal of a proper fraction or mixed number.
227	multiply and/or divide a mixed number by a mixed number.
228	solve word problems of a consumer and/or career nature involving multiplication and division of fractions and mixed numbers.
230	find the lowest common denominator of two or three mixed numbers.
233	add two mixed numbers with like or unlike denominators.
234	subtract two mixed numbers with like or unlike demoninators and regroup.
235	solve word problems of a consumer and/or career nature involving addition and subtraction of fractions or mixed numbers.

Level 3

Code	Objective: "The student will be able to. . ."
302	find the probabilities using tree diagrams.
305	use experiments to make predictions.
306	solve word problems involving probability.
309	find the missing number in a proportion by using cross products.
310	interpret and use scale drawings and maps, when the scale factor has been given, to determine appropriate distance.
311	identify pairs of congruent segments or polygons, recognize a pair of lines which are parallel or a pair which are perpendicular and construct a perpendicular bisector of a segment.
312	identify acute, right, obtuse and congruent angles.

Ninth Grade General Mathematics I Objectives
Assessed by the GMI Area Exam

Level 3 (continued)

Code	Objective: "The student will be able to. . ."
313	identify triangle, qualrilateral, pentagon, hexagon, octagon, prism, cube, pyramid, sphere, cylinder and cones.
314	identify similar figures.
315	identify corresponding angles and sides of similar figures.
316	find missing dimensions in similar triangles by using proportion.
317	solve word problems involving ratio, proportion and similarity.
318	write percents as decimals and vice versa.
319	write percents as fractions and vice versa.
320	find a percent of a number.
321	find what percent one number is of another.
322	find a number when a percent of it is known.
323	solve word problems of a consumer and/or career nature involving percent (include interest problems).
324	find the mean of a set of numbers.
325	find the median and the mode of a set of numbers.
326	read and make a bar graph.
327	read and make a line graph.
328	read and make a circle graph.
329	solve word problems of a consumer and/or career nature involving statistics (mean, median, mode, graphs).
330	plot positive and negative numbers on a number line.
331	select the proper inequality or equality sign to compare positive and negative numbers.
332	arrange a set of integers in order from least to greatest.
333	add positive and negative numbers.
334	subtract positive and negative numbers.

Ninth Grade General Mathematics I Objectives
Assessed by the GMI Area Exam

Level 3 (continued)

Code	Objective: "The student will be able to. . ."
336	divide positive and negative numbers.
337	solve word problems of a consumer and/or career nature involving positive and negative numbers.

Level 4

Code	Objective: "The student will be able to. . ."
402	evaluate expressions involving order of operations. (involving variables)
403	solve addition and subtraction equations involving integers and decimals.
404	solve multiplication and division equations involving integers and decimals.
405	solve two-step equations.
406	combine like terms to solve equations.
407	solve word problems of a consumer and/or career nature involving equations.
411	give ordered pairs of integers for points in four quadrants and locate points for ordered pairs of integers.
412	read the graph of an equation on a four quadrant grid to complete a table of ordered pairs.
413	draw the graph of a given equation on a four quadrant grid.
414	solve word problems of a consumer and/or career nature involving coordinate graphing.
415	find the perimeter of a geometric figure.
416	find the circumference of a circle.
417	find the area of a rectangle, a square and a parellelogram.
418	find the area of a triangle and a trapezoid.
419	find the area of a circle, or find the radius and give the area.

Ninth Grade General Mathematics I Objectives
Assessed by the GMI Area Exam

Level 4 (continued)

Code	Objective: "The student will be able to. . ."
420	solve word problems of a consumer and/or career nature involving perimeter and area.
424	find the volume of a rectangular prism and a cube.
425	find the volume of a cylinder.
426	find the volume of a pyramid and a cone.
427	find the volume of a sphere.
428	solve word problems of a consumer and/or career nature involving surface area and volume.
430	measures angles by using a protractor.